On Similarity Invariance of Balancing for Nonlinear Systems

Scherpen, Jacquelien M.A.

Published in:
Preprints of the Nonlinear Control Systems Design Symposium NOLCOS '95

On Similarity Invariance of Balancing for Nonlinear Systems

Jacquelien M.A. Scherpen*

* *Department of Electrical Engineering, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands. E-mail: J.M.A.Scherpen@et.tudelft.nl*

Abstract. A previously obtained balancing method for nonlinear systems is investigated on similarity invariance by generalization of the observations on the similarity invariance of the linear balanced realization theory. For linear systems it is well known that the Hankel singular values are similarity invariants. It is shown that under some additional conditions a similar statement on the similarity invariance holds for nonlinear systems. To be able to do so the concepts of local zero-state observability and local strong accessibility are considered. The local strong accessibility is an additional condition, which is needed to consider a nonlinear generalization of the Kalman decomposition. In the local coordinates that correspond to this decomposition the controllability and observability functions are investigated.

Key Words. balancing, Hamilton-Jacobi equations, singular value functions, similarity invariance.

1. INTRODUCTION

Balancing for stable linear systems has been introduced by (Moore, 1981), and turned out to be a useful tool to analyze a linear system, and to apply it to model reduction. In the balancing method for stable linear systems the Hankel singular values play an important role, and have the nice property that they are similarity invariants, i.e., independent of the chosen state space realization, and thus only dependent of the input-output behavior of the system. Since its introduction, the balancing theory for linear systems has been explored further into several directions, e.g. (Jonckheere and Silverman, 1983), (Ober and McFarlane, 1989), (Meyer, 1990), (Mustafa and Glover, 1991). All of the obtained balancing methods have the same property with respect to similarity invariance.

Balancing for stable nonlinear systems has been introduced recently, see (Scherpen, 1993*a*), (Scherpen, 1993*b*), (Scherpen, 1994), and deals with the past input and the future output energy functions of the system. The singular value functions of the nonlinear system are obtained from these energy functions, and they equal the squared Hankel singular values in case of a linear system.

In this paper we investigate to what extend the observations on the similarity invariance for the linear balancing theory can be generalized to the nonlinear balancing theory. To bring a stable nonlinear system in balanced form, we need the system to fulfill conditions on zero-observability and anti-stabilizability. In the linear case the condition on anti-stabilizability is equivalent to controllability, which is not true in the nonlinear case. To deal with this difference we additionally make assumptions on accessibility of the nonlinear system. Then we use the nonlinear generalization of the Kalman decomposition to investigate similarity invariance of the nonlinear balancing theory.

In Section 2 we give a review on balancing for stable nonlinear systems. Section 3 contains the analysis of the similarity invariance of this method. Finally, in Section 4 we give the conclusions.

Throughout this paper we will use a fairly standard notation. We denote by $x^T x$ or $\| x \|^2$ the squared norm of a vector $x \in \mathbb{R}^n$. We say that $u : (-\infty, 0) \to \mathbb{R}^m$ is in $L_2(-\infty, 0)$ if $\int_{-\infty}^{0} \| u(t) \|^2 \, dt < \infty$. By $\frac{\partial L}{\partial x}(x)$ we denote the row-vector of partial derivatives of a differentiable function $L : \mathbb{R}^n \to \mathbb{R}$. Furthermore we denote by $x(t_2) = \varphi(t_2, t_1, x_1, u)$ the solution on time t_2 of the system $\dot{x} = f(x) + g(x)u$ with initial condition $x(t_1) = x_1$ and input $u : [t_1, t_2] \to \mathbb{R}^m$.

2. REVIEW ON BALANCING FOR STABLE NONLINEAR SYSTEMS

Balancing for stable nonlinear systems is dealt with in (Scherpen, 1993*a*). As in the linear case, this is a method based on the input energy that is necessary to reach a state and the output energy that is generated by this state. We will give a brief review on this subject.

Consider a smooth, i.e., C^∞, nonlinear system of the form

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x) \end{aligned} \tag{1}$$

where $u = (u_1, \ldots, u_m) \in \mathbb{R}^m$, $y = (y_1, \ldots, y_p) \in \mathbb{R}^p$ and $x = (x_1, \ldots, x_n)$ are local coordinates for a smooth state space manifold denoted by M. Throughout we assume that the system has an equilibrium. Without loss of generality we take this equilibrium in 0, i.e. $f(0) = 0$ and we also take $h(0) = 0$.

Definition 1 The controllability and observability function of a nonlinear system are defined as

$$L_c(x_0) = \min_{\substack{u \in L_2(-\infty,0) \\ x(-\infty)=0, x(0)=x_0}} \frac{1}{2}\int_{-\infty}^{0} ||u(t)||^2 dt \tag{2}$$

and

$$L_o(x_0) = \frac{1}{2}\int_0^{\infty} ||y(t)||^2 dt, \tag{3}$$

$$x(0) = x_0, \qquad u(t) \equiv 0, \ \ 0 \leq t < \infty,$$

respectively. □

The value of the controllability function at x_0 is the minimum amount of control energy required to reach the state x_0 and the value of the observability function at x_0 is the amount of output energy generated by x_0. We throughout assume L_c and L_o are *finite*. Also, for the rest of this paper we assume L_c and L_o are *smooth* functions of x.

Theorem 1 (Scherpen, 1993*a*) If $f(x)$ is asymptotically stable on a neighborhood W of 0, then for all $x \in W$, $L_o(x)$ is the unique smooth solution of the following Lyapunov type of equation:

$$\frac{\partial L_o}{\partial x}(x) f(x) + \frac{1}{2} h^T(x) h(x) = 0, \tag{4}$$

$L_o(0) = 0$. Furthermore for all $x \in W$, $L_c(x)$ is the unique smooth solution of the following Hamilton-Jacobi equation:

$$\frac{\partial L_c}{\partial x}(x) f(x) + \frac{1}{2}\frac{\partial L_c}{\partial x}(x) g(x) g^T(x) \frac{\partial^T L_c}{\partial x}(x) = 0,$$

$$L_c(0) = 0 \tag{5}$$

satisfying $-(f(x) + g(x)g^T(x)\frac{\partial^T L_c}{\partial x}(x))$ is asymptotically stable on W. □

Remark 1 (Scherpen, 1993*a*) L_c and L_o are *non-negative*. □

Remark 2 If we assume that $f(x)$ is asymptotically stable and that (4) has a smooth solution, it follows that L_o, as in (3), exists, i.e. is finite. Furthermore, if we assume that (5) has a smooth solution L and that $-(f(x) + g(x)g(x)^T \frac{\partial^T L}{\partial x}(x))$ is asymptotically stable, it follows that L_c, as in (2), exist, i.e. is finite. □

Theorem 2 (Scherpen, 1993*a*) Assume f is asymptotically stable on W and (5) has a smooth solution $\bar{L}_c$ on W. Then $\bar{L}_c(x_0) > 0$ for $x_0 \in W$, $x_0 \neq 0$, if and only if $-(f(x) + g(x)g^T(x)\frac{\partial^T \bar{L}_c}{\partial x}(x))$ is asymptotically stable on W. □

For the following definitions see e.g. (Hill and Moylan, 1976), (van der Schaft, 1992), (Nijmeijer and van der Schaft, 1990).

Definition 2 The system (1) is *reachable from* x_0 if for any $\bar{x} \in M$ there exists a $\bar{t} \geq 0$ and input u such that $\bar{x} = \varphi(\bar{t}, 0, x_0, u)$.
The system is *locally strongly accessible* from x_0 if for any neighborhood V of x_0 the set $R^V(x_0, T)$ (the reachable set from x_0 at time $T > 0$, following the trajectories which remain in the neighborhood V of x_0 for $t \leq T$) contains a non-empty open set for any $T > 0$ sufficiently small.
The system (1) is *zero-state observable* if any trajectory where $u(t) \equiv 0, y(t) \equiv 0$ implies $x(t) \equiv 0$, i.e., for all $x \in M$, $h(\varphi(t,0,x,0)) = 0, t \geq 0 \Rightarrow \varphi(t,0,x,0) = 0, t \geq 0$.
The system (1) is *locally zero-state observable* at 0, if there exists a neighborhood W of 0 such that for all $x \in W$, $h(\phi(t,0,x,0)) = 0$, for all $t \geq 0 \Rightarrow \varphi(t,0,x,0) = 0$ for all $t \geq 0$. □

For local zero-state observability we can give a condition in terms of Lie derivatives. This is closely related to the condition for local observability in terms of the observability codistribution, e.g. (Nijmeijer and van der Schaft, 1990). Furthermore, we give a condition in terms of Lie brackets for local strong accessibility, which is well known, e.g. (Nijmeijer and van der Schaft, 1990).

Definition 3 Consider the nonlinear system (1). The *strong accessibility algebra* $\mathcal{C}_0$ is the smallest subalgebra of $V^\infty(M)$ (the Lie algebra of vector fields on M) that contains $g_1, \cdots, g_m$ and satisfies $[f, X] \in \mathcal{C}_0$ for all $X \in \mathcal{C}_0$.
The *strong accessibility distribution* C_0 is the distribution generated by $\mathcal{C}_0$, i.e., $C_0(x) = \text{span}\{X(x) \,|\, X \text{ vector field in } \mathcal{C}_0\}$.
Consider the nonlinear system (1). The *zero-observation space* $\mathcal{O}_0$ of (1) is the linear space of functions on M containing $h_1, \ldots, h_p$ and all repeated Lie derivatives $L_f^k h_j$, $j \in 1, \ldots, p$, $k = 1, 2, \ldots$
The *zero-observability codistribution* $d\mathcal{O}_0$ is given by $d\mathcal{O}_0(q) = \text{span}\{dH(q) \,|\, H \in \mathcal{O}_0\}$, where $q \in M$. □

Theorem 3 Consider the system (1). If $\dim C(x_0) = n$, then de system is locally strongly accessible from 0. If $\dim d\mathcal{O}_0(0) = n$, then the system is locally zero-state observable at 0.

Proof: The proof is well-known/follows well-known arguments as may be found in (Nijmeijer and van der Schaft, 1990). ■

The following theorem is closely related to some results in (Hill and Moylan, 1976) and (van der Schaft, 1992). For the proof, see (Scherpen, 1993*a*).

Theorem 4 Assume $f(x)$ is asymptotically stable on a neighborhood W of 0. If the system (1) is zero-state observable on W, then $L_o(x_0) > 0, \forall x_0 \in W, x_0 \neq 0$. □

Now we consider nonlinear systems of the form (1) with controllability and observability function L_c respectively L_o as in Definition 1, and with the following additional assumptions:

1. $f(x)$ is asymptotically stable on some neighborhood Y of 0
2. $L_c(x)$ and $L_o(x)$ are smooth and finite functions of x on Y
3. the system is zero-state observable on Y
4. $\frac{\partial^2 L_c}{\partial x^2}(0) > 0$ and $\frac{\partial^2 L_o}{\partial x^2}(0) > 0$

Lemma 1 (Scherpen, 1993*a*) There exists a coordinate transformation $x = \phi(\bar{x})$, $\phi(0) = 0$, such that $L_c(x)$ in the new coordinates $\bar{x} = \phi^{-1}(x)$ is of the following form:

$$L_c(\phi(\bar{x})) = \frac{1}{2}\bar{x}^T\bar{x} \tag{6}$$

Furthermore we can write $L_o(x)$ in the new coordinates $\bar{x} = \phi^{-1}(x)$ in the following form:

$$L_o(\phi(\bar{x})) = \frac{1}{2}\bar{x}^T M(\bar{x})\bar{x} \tag{7}$$

$$\text{where} \quad M(0) = \frac{\partial^2 L_o}{\partial x^2}(0)$$

with $M(\bar{x})$ a $n \times n$ symmetric matrix with entries which are smooth functions of $\bar{x}$. □

Comparing with the linear situation we see that we are close to an input-normal form if we can bring $M(\bar{x})$ into a diagonal form, while we keep the form of the controllability function as above. To be able to do so, we need the following lemma.

Lemma 2 (Kato, 1982),(Scherpen, 1993*a*) If there exists a neighborhood V of 0 where the number of distinct eigenvalues of $M(\bar{x})$ is constant for $\bar{x} \in V$, then on V the eigenvalues $\lambda_i(\bar{x})$, $i = 1,\ldots,n$, are smooth functions of $\bar{x}$, as well as the associated normalized eigenvectors. □

Theorem 5 (Scherpen, 1993*a*) Consider system (1) and assume there exists a neighborhood V of 0 where the number of distinct eigenvalues of $M(\bar{x})$ is constant for $\bar{x} \in V$. Then there exists a neighborhood U of zero and a coordinate transformation $x = \psi(z)$, $\psi(0) = 0$, such that $L_c(x)$ in the new coordinates $z \in W := \psi^{-1}(U)$ is of the following form:

$$\tilde{L}_c(z) := L_c(\psi(z)) = \frac{1}{2}z^T z \tag{8}$$

while in the new coordinates $\tilde{L}_o(z) := L_o(\psi(z))$ is of the form

$$\tilde{L}_o(z) = \frac{1}{2}z^T \begin{pmatrix} \tau_1(z) & & 0 \\ & \ddots & \\ 0 & & \tau_n(z) \end{pmatrix} z \tag{9}$$

where $\tau_1(z) \geq \ldots \geq \tau_n(z)$ are smooth functions of z, called the singular value functions. □

Remark 3 For a linear system the singular value functions τ_i, $i = 1,..,n$ are constant and are equal to the squared Hankel singular values. □

The form of the controllability and observability function in (8) and (9) is not yet entirely balanced. For that we need another additional coordinate transformation. We take as smooth transformation $\bar{z}_i = \eta_i(z_i) := \tau_i(0,..,0,z_i,0,..0)^{\frac{1}{4}} z_i$, $i = 1,..,n$ and hence $\bar{z} = \eta(z) := (\eta_1(z_1) \ \ldots \ \eta_n(z_n))$ on $\bar{z} \in \bar{W} := \eta(W)$. Define $\hat{L}_c(\bar{z}) := \tilde{L}_c(\eta^{-1}(\bar{z}))$ and $\hat{L}_o(\bar{z}) := \tilde{L}_o(\eta^{-1}(\bar{z}))$. Then (8) and (9) become respectively:

$$\hat{L}_c(\bar{z}) = \frac{1}{2}\bar{z}^T \begin{pmatrix} \sigma_1(\bar{z}_1)^{-1} & & 0 \\ & \ddots & \\ 0 & & \sigma_n(\bar{z}_n)^{-1} \end{pmatrix} \bar{z} \tag{10}$$

$$\hat{L}_o(\bar{z}) = \frac{1}{2}\bar{z}^T G(\bar{z})\bar{z} \quad \text{where} \quad G(\bar{z}) = \tag{11}$$

$$\begin{pmatrix} \sigma_1(\bar{z}_1)^{-1}\tau_1(\eta^{-1}(\bar{z})) & & 0 \\ & \ddots & \\ 0 & & \sigma_n(\bar{z}_n)^{-1}\tau_n(\eta^{-1}(\bar{z})) \end{pmatrix}$$

and where $\sigma_i(\bar{z}_i) = \tau_i(0,..,0,\eta_i^{-1}(\bar{z}_i),0,..,0)^{\frac{1}{2}}$ for $i = 1,..,n$. It follows that $\hat{L}_c(0,..,0,\bar{z}_i,0,..,0) = \frac{1}{2}\bar{z}_i^2\sigma_i(\bar{z}_i)^{-1}$ and $\hat{L}_o(0,..,0,\bar{z}_i,0,..,0) = \frac{1}{2}\bar{z}_i^2\sigma_i(\bar{z}_i)$ for $i = 1,..,n$. We call a nonlinear system *balanced* if it has a controllability and observability function of the form of respectively (10) and (11). This means that we can balance system (1) by a coordinate transformation of the form $x = \chi(\bar{z}) := \psi(\eta^{-1}(\bar{z}))$ for $\bar{z} \in \bar{W}$, where ψ is given in Theorem 5.

3. SIMILARITY INVARIANTS

For linear systems it is well known that the Hankel singular values are similarity invariants, i.e., the Hankel singular values are independent of the chosen state space realization and only depend on the input-output behavior of the system. In fact, they are the singular values of the Hankel operator of the system, e.g. (Glover, 1984). If we consider a linear state space system that is not minimal, and we study its controllability Gramian W and observability Gramian M, then the non-zero eigenvalues of MW equal the squared Hankel singular values of its input-outputmap, and the number of zero eigenvalues of MW equals the difference of the state space dimension of the non-minimal system and the state space dimension of a minimal representation of it.

Furthermore, for linear systems a balanced representation is almost *unique* in the following sense. Assume that there are k distinct Hankel singular values, and that the singular value σ_i has multiplicity j_i, $i = 1,\ldots,n$. The balanced realization is unique up to linear transformations of the form

$$\begin{pmatrix} T_1 & & 0 \\ & \ddots & \\ 0 & & T_k \end{pmatrix}, \tag{12}$$

where the blocks T_i, $i = 1,\ldots,k$, are $j_i \times j_i$ orthogonal matrices, i.e., $T_i^T T_i = I$, e.g. (Glover, 1984).

In this section we want to extend these observations to the nonlinear case. We consider the nonlinear system (1) and we assume that it is locally asymptotically stable. However, we do *not* assume local zero-state observability, and hence the observability function is not necessarily positive definite. Furthermore, we do *not* assume that $-(f(x)+g(x)g(x)^T \frac{\partial^T L_c}{\partial x}(x))$ is locally asymptotically stable, and thus the controllability function need not be finite for all x. (Recall that the controllability and observability function, L_c and L_o, respectively, are defined as in Definition 1.)

We use Frobenius' Theorem, see e.g. (Nijmeijer and van der Schaft, 1990), to construct the zero-state observable 'part' of the system.

Theorem 6 Assume that $d\mathcal{O}_0$ has constant dimension k. By Frobenius' Theorem we may find local coordinates $(x^1,x^2) = (x_1,\ldots,x_k,x_{k+1},\ldots,x_n)$ such that $\ker d\mathcal{O}_0 = \{\frac{\partial}{\partial x^2}\}$. Write correspondingly $f = (f^1,f^2)^T$ and $g = (g^1,g^2)^T$. In these coordinates the system (1) takes the form

$$\begin{aligned} \dot{x}^1 &= f^1(x^1)+g^1(x^1,x^2)u & (13)\\ \dot{x}^2 &= f^2(x^1,x^2)+g^2(x^1,x^2)u & (14)\\ y &= h(x^1) & (15) \end{aligned}$$

Proof: The codistribution $d\mathcal{O}_0$ is invariant for the dynamics $\dot{x} = f(x)$, since $L_f d\mathcal{O}_0 \subset d\mathcal{O}_0$. Then $\ker d\mathcal{O}_0 = \{\frac{\partial}{\partial x^2}\}$ is an invariant distribution for $\dot{x} = f(x)$. Since $\ker d\mathcal{O}_0 \subset \ker dh$, (1) takes the form (13), (14) and (15) (see Proposition 3.42 in (Nijmeijer and van der Schaft, 1990)). ∎

Now we also consider the part of the state space system where

$$-\left(f(x)+g(x)g^T(x)\frac{\partial L_c}{\partial x}(x)\right) \tag{16}$$

is asymptotically stable. In the linear case this part is equal to the controllable part of the system, since then the asymptotic stability of (16) is equivalent with controllability of the system. (Recall that f is assumed to be asymptotically stable.) For nonlinear systems this is not always the case, but in order to be able to construct a decomposition as in the nonlinear generalization of the Kalman decomposition (e.g., Theorem 3.51 in (Nijmeijer and van der Schaft, 1990)), we make an additional assumption. That is, we assume that similarly to the linear case *the part of the system where (16) is asymptotically stable equals the strongly accessible part of the system* (Definition 2).

Theorem 7 Assume that the distributions C_0, $\ker d\mathcal{O}_0$ and $C_0 + \ker d\mathcal{O}_0$ all have constant dimension and that $C_0 + \ker d\mathcal{O}_0$ is involutive. Then we can find local coordinates $x = (x^1,x^2,x^3,x^4)$ such that $C_0 = \operatorname{span}\{\frac{\partial}{\partial x^1},\frac{\partial}{\partial x^2}\}$ and $\ker d\mathcal{O}_0 = \operatorname{span}\{\frac{\partial}{\partial x^2},\frac{\partial}{\partial x^4}\}$. The system takes the form

$$\begin{aligned} \dot{x}^1 &= f^1(x^1,x^3)+\sum_{j=1}^m g_j^1(x^1,x^2,x^3,x^4)u_j & (17)\\ \dot{x}^2 &= f^2(x^1,x^2,x^3,x^4)+\sum_{j=1}^m g_j^2(x^1,x^2,x^3,x^4)u_j & (18)\\ \dot{x}^3 &= f^3(x^3) & (19)\\ \dot{x}^4 &= f^4(x^3,x^4) & (20)\\ y &= h(x^1,x^3) & (21) \end{aligned}$$

Proof: We may apply an extension of Frobenius' Theorem. Then as in Theorem 6 the form follows. ∎

Let n_i be the dimension of x^i, $i = 1,2,3,4$, and let Y be a neighborhood of 0 such that the decomposition as above can be done for $x \in Y$. Then clearly (17), (19) and (21) form the zero-state observable part of the system, while (17) and (18) is the strongly accessible part of the system. To assure that for (17), (19) and (21) the observability function exists, we assume that in these local coordinates the following equation

$$\begin{aligned} &\frac{\partial \check{L}_o}{\partial x^1}(x^1,x^3)f^1(x^1,x^3)+\frac{\partial \check{L}_o}{\partial x^3}(x^1,x^3)f^3(x^3)+ \\ &\frac{1}{2}h(x^1,x^3)^T h(x^1,x^3) = 0, \end{aligned} \tag{22}$$

$\check{L}_o(0) = 0$, has a smooth solution for $(x^1,0,x^3,0) \in Y$. Furthermore, note that

$$\begin{pmatrix} f^3(x^3) \\ f^4(x^3,x^4) \end{pmatrix}$$

is asymptotically stable and by the form of (19), and

(20) it is impossible that $-(f(x)+g(x)g(x)^T \frac{\partial^T L_c}{\partial x}(x))$ is asymptotically stable on Y. To assure that for (17) and (18) the controllability function exists, we assume that in these local coordinates the following equation

$$\frac{\partial \check{L}_c}{\partial x^1}(x^1,x^2)f^1(x^1,0)+\frac{\partial \check{L}_c}{\partial x^2}(x^1,x^2)f^2(x^1,x^2,0,0)+$$

$$\frac{1}{2}\left(\frac{\partial \check{L}_c}{\partial x^1}(x^1,x^2) \quad \frac{\partial \check{L}_c}{\partial x^2}(x^1,x^2)\right)\cdot\begin{pmatrix} g^1(x^1,x^2,0,0)\\ g^2(x^1,x^2,0,0)\end{pmatrix}\cdot \tag{23}$$

$$\left(g^1(x^1,x^2,0,0)^T \quad g^2(x^1,x^2,0,0)^T\right)\cdot\begin{pmatrix}\frac{\partial^T \check{L}_c}{\partial x^1}(x^1,x^2)\\ \frac{\partial^T \check{L}_c}{\partial x^2}(x^1,x^2)\end{pmatrix}$$

$$=0$$

with $\check{L}_c(0)=0$, has a smooth solution for $(x^1,x^2,0,0)\in Y$, such that

$$-[\begin{pmatrix} f^1(x^1,0)\\ f^2(x^1,x^2,0,0)\end{pmatrix}+\begin{pmatrix} g^1(x^1,x^2,0,0)\\ g^2(x^1,x^2,0,0)\end{pmatrix}\cdot \tag{24}$$

$$\left(g^1(x^1,x^2,0,0)^T \quad g^2(x^1,x^2,0,0)\right)\cdot\begin{pmatrix}\frac{\partial^T \check{L}_c}{\partial x^1}(x^1,x^2)\\ \frac{\partial^T \check{L}_c}{\partial x^2}(x^1,x^2)\end{pmatrix}]$$

is asymptotically stable for $(x^1,x^2,0,0)\in Y$.

Theorem 8 If the assumptions above are fulfilled, then $L_o(x^1,x^2,x^3,x^4)>0$ whenever $(x^1,x^4,\ x^3,x^4)\in Y$, and $(x^1,x^3)\neq(0,0)$, and $L_o(0,x^2,0,x^4)=0$ for all $(0,x^2,0,x^4)\in Y$.
Furthermore, $L_c(x^1,x^2,x^3,x^4)$ is infinite whenever $(x^1,x^2,x^3,x^4)\in Y$, $(x^3,x^4)\neq(0,0)$, and $0<L_c(x^1,x^2,0,0)<\infty$, for all $(x^1,x^2,0,0)\in Y$, $(x^1,x^2)\neq(0,0)$.

Proof: It is clear that $h(0,x^2(\tau),0,x^4(\tau))=0$ for all $\tau\geq 0$, and by the form of (17) and (19) we obtain that

$$L_o(0,x^2,0,x^4)=$$

$$\frac{1}{2}\int_0^\infty h(0,x^2(\tau),0,x^4(\tau))^T h(0,x^2(\tau),0,x^4(\tau))d\tau$$

$$=0,$$

for $u\equiv 0$, and for all $(0,x^2,0,x^4)\in Y$. Again, by the form of (17), (19), (21) we have

$$L_o(x^1,x^2,x^3,x^4)=$$

$$\frac{1}{2}\int_0^\infty h(x^1(\tau),x^3(\tau))^T h(x^1(\tau),x^3(\tau))d\tau$$

$$=\check{L}_o(x^1,x^3),$$

for $u\equiv 0$, where $\check{L}_o$ is the observability function of (17), (19), (21). We assumed that (22) has a smooth solution and thus, $\check{L}_o=L_o$ exists and is smooth and by Theorem 4 $L_o(x^1,x^2,x^3,x^4)=\check{L}_o(x^1,x^3)>0$ for $(x^1,x^3)\neq(0,0)$.
The controllability function L_c is

$$L_c(x_0)=\min_{\substack{u\in L_2(-\infty,0)\\ x(-\infty)=0,x(0)=x_0}}\frac{1}{2}\int_{-\infty}^0 \| u(t)\|^2\,dt.$$

Since

$$\begin{pmatrix} f^3(x^3)\\ f^4(x^3,x^4)\end{pmatrix}$$

is asymptotically stable, it follows immediately that $L_c(x^1,x^2,x^3,x^4)=\infty$ for all $(x^1,x^2,x^3,x^4)\in Y$, with $(x^3,x^4)\neq(0,0)$. By (23), the asymptotic stability of (24), and Theorem 1 it follows that $L_c(x^1,x^2,0,0)=\check{L}_c(x^1,x^2)<\infty$ for all $(x^1,x^2,0,0)\in Y$. Furthermore, by Theorem 2 it follows that $L_c(x^1,x^2,0,0)>0$ for all $(x^1,x^2,0,0)\in Y$, $(x^1,x^2)\neq(0,0)$. ∎

If we additionally assume that

- $\frac{\partial^2 L_o}{\partial x^1\partial x^1}(0)>0$ and $\frac{\partial^2 L_c}{\partial x^1\partial x^1}(0)>0$,

it becomes clear from this theorem that $L_o(x^1,0,0,0)$ and $L_c(x^1,0,0,0)$ may be transformed to the form of Theorem 5, if the condition of Lemma 2 is fulfilled. In fact there exists a local x^1 coordinate transformation $x^1=\psi(z)$, $\psi(0)=0$, $(\psi^{-1}(x^1),0,0,0)\in Y$, such that

$$L_c(\psi(z),0,0,0)=\frac{1}{2}z^Tz$$

and

$$L_o(\psi(z),0,0,0)=\frac{1}{2}z^T\begin{pmatrix}\tau_1(z) & & 0\\ & \ddots & \\ 0 & & \tau_{n_1}(z)\end{pmatrix}z.$$

Thus this part of the system may be balanced on a neighborhood of 0, with singular value functions $\tau_1(z)\geq\cdots\geq\tau_{n_1}(z)$.

Furthermore, if we also consider x^2, then there exist local coordinates $(z^1,z^2)=\phi^{-1}(x^1,x^2)$ such that

$$L_c(\phi(z^1,z^2),0,0)=\frac{1}{2}\left(z^{1T}\quad z^{2T}\right)\begin{pmatrix}z^1\\ z^2\end{pmatrix}.$$

Now we may apply Lemma 1 and Lemma 2 to the observability function. Write $L_o(\phi(z^1,z^2),0,0)$ in the form of Lemma 1, i.e.,

$$L_o(\phi(z^1,z^2),0,0)=\frac{1}{2}\left(z^{1T}\quad z^{2T}\right)M(z^1,z^2)\begin{pmatrix}z^1\\ z^2\end{pmatrix}.$$

If the condition of Lemma 2 is fulfilled, we may diagonalize $M(z^1,z^2)$. Then we will find as functions on the diagonal $\bar{\tau}_1(z^1,z^2)\geq\cdots\geq\bar{\tau}_{n_1+n_2}(z^1,z^2)$, where $\bar{\tau}_i(z^1,0)=\tau_i(z)$, $i=1,\ldots,n_1$, and $\bar{\tau}_j(0,x^2)=0$, $j=n_1+1,\ldots,n_1+n_2$. This is in accordance with the linear case, where the unobservable part gives zero 'Hankel singular values'. Note that it is not possible to transform the whole system to the form of Theorem 5, since $L_c(0,0,x^3,x^4)$ is infinite. This is still in accordance with the linear theory, since here we are dealing with the 'inverse of the controllability Gramian'. Hence the part of the system that is not strongly accessible yields an 'inverse of the controllability Gramian' that is infinite, and thus a 'controllability Gramian' that is zero.

In order to study the uniqueness of balanced representations of nonlinear systems, let us consider two balanced representations of one nonlinear system, satisfying all conditions of Section 2. Clearly, both representations are linked via a coordinate transformation. Let k be the number of distinct singular value functions, and let j_i be the number of times the i^{th} singular value function appears. Then it follows directly that both systems have the same balanced form, except for a coordinate transformation of a similar form as in the linear case (transformations of the form (12)). The singular value functions belonging to these forms are related by the same coordinate transformation. In the nonlinear case the transformations are of the form

$$x = \begin{pmatrix} T_1(\bar{x}) & & 0 \\ & \ddots & \\ 0 & & T_k(\bar{x}) \end{pmatrix} \bar{x} \tag{25}$$

where the blocks $T_i(\bar{x})$, $i = 1,\ldots,k$, are $j_i \times j_i$ orthogonal matrices, i.e., $T_i(\bar{x})^T T_i(\bar{x}) = I$, with entries that are smooth functions of $\bar{x}$. This easily follows from the form of the controllability function (8) and observability function (9) in Theorem 5.

4. CONCLUSIONS

We investigated the similarity invariance of the nonlinear balancing method. The analysis in this paper leads to the conclusion that under some assumptions the singular value functions are, as in the linear case, 'similarity invariants', i.e., independent of the state space representation (minimal or not). From this conclusion it follows that it is natural to consider the zero-state observable and strongly accessible part, i.e., the 'minimal' part of the system for analyzing the controllability and observability functions. Furthermore, the singular value functions of the zero-state observable and strongly accessible part are invariant except for a coordinate transformation of the form (25). Thus, the singular value functions as functions of the state may change by such a (quite restricted) coordinate transformation, but it leaves the the singular value functions as functions of the time t invariant.

ACKNOWLEDGEMENT

This research has been supported by the Dutch Systems and Control Theory Network.
I like to thank Arjan van der Schaft for the useful discussions and his advice on this research.

5. REFERENCES

Glover, K. (1984). All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds. *Int. J. of Contr.* **39**, 1115–1193.

Hill, D. and P. Moylan (1976). The stability of nonlinear dissipative systems. *IEEE Trans. Aut. Contr.* **AC-21**, 708–711.

Jonckheere, E.A. and L.M. Silverman (1983). A new set of invariants for linear systems - applications to reduced order compensator design. *IEEE Trans. Aut. Cont.* **AC-28**, 953–964.

Kato, T. (1982). *A Short Introduction to Perturbation Theory for Linear Operators*. Springer-Verlag, New York.

Meyer, D.G. (1990). Fractional balanced reduction - model reduction via a fractional representation. *IEEE Trans. Aut. Contr.* **AC-26**, 1341–1345.

Moore, B.C. (1981). Principal component analysis in linear systems: Controllability, observability and model reduction. *IEEE Trans. Aut. Contr.* **AC-26**, 17–32.

Mustafa, D. and K. Glover (1991). Controller reduction by $\mathcal{H}_\infty$-balanced truncation. *IEEE Trans. Aut. Contr.* **AC-36**, 668–682.

Nijmeijer, H. and A.J. van der Schaft (1990). *Nonlinear Dynamical Control Systems*. Springer-Verlag, New York.

Ober, R. and D. McFarlane (1989). Balanced canonical forms for minimal systems: A normalized coprime factor approach. *Linear Algebra and its Applications* pp. 23–64.

van der Schaft, A.J. (1992). L_2-gain analysis of nonlinear systems and nonlinear state feedback $\mathcal{H}_\infty$ control. *IEEE Trans. Aut. Contr.* **AC-37**, 770–784.

Scherpen, J.M.A. (1993*a*). Balancing for nonlinear systems. *Syst. & Contr. Letters* **21**, 143–153.

Scherpen, J.M.A. (1993*b*). Balancing for nonlinear systems. In: *Eur. Contr. Conf.*. Vol. 4. pp. 1838–1843.

Scherpen, J.M.A. (1994). Balancing for nonlinear systems. PhD thesis. University of Twente.